

PanDA Motivation: History & Lessons Learned

Kaushik De
University of Texas at Arlington

ePIC S&C Weekly, 11 March 2026
<https://indico.bnl.gov/event/31460/>

The starting point

- About 25 years ago
 - HEP experiments were becoming global at the LHC
 - But HEP computing facilities remained mostly local
 - Experiments provided primary computing facility at host lab
 - A few satellite facilities were used for CPU intensive simulations
 - Physics analysis was done on batch systems at universities
- ATLAS computing challenge at the LHC
 - CERN was not going to host majority of computing or data
 - 90% of the computing cycles and data storage were widely distributed at collaborating labs and universities
 - Physics analyzers were based worldwide at home institutes
- Additional concern for US collaborators
 - Small computing support team in the US, far from CERN

The First Attempt

- ATLAS Data Challenge 0 (DC0) was held in this milieu
 - About 30-40 computing centers worldwide participated
 - Huge coordinated global exercise, with local execution of tasks
 - A 10% scale exercise to test a new computing model for the LHC
 - Use Tevatron and LEP experience but distributed along Tiers
- Lessons from DC0
 - Enormous team needed to execute computing globally
 - Time consuming exercise requiring complex coordination
 - But we did it differently in the US – we distributed the production among ~5 sites using the (new) Globus grid toolkit
 - US experience with grid showed a new path forward
 - System developed and run by a small team, led by UTA (De)
- ATLAS decided to develop a grid solution for next DC1

The Second Attempt

- The new distributed computing system was based on 2 pillars
 - Automated distributed data processing system
 - Separate data catalogue and data movement system
- Lessons from DC1
 - Tested at ~10% scale with MC simulations
 - Using the grid was a huge success – no manual processing
 - Data Processing software team led by UTA (De) and CERN (Goosens), Data Management system led by BNL (Wenaus)
 - Software based on messaging queues and distributed agents
 - But site and network outages required frequent and daily manual interventions – heavy operations load
 - Production system was fragile due to unexpected errors
 - Lack of full integration between two systems
 - User analysis still available to only a few people at few sites

Third and Last - PanDA

- A new system was developed for DC2 in 2005 – PanDA
 - PanDA – **P**roduction **and** **D**istributed **A**nalysis system
 - Led by UTA (De) and BNL (Wenaus), initially used in the US only
 - Later adopted ATLAS-wide before the start of Run 1 at the LHC
 - Also included distributed analysis within the same system
- Goals
 - **Improve user experience** by making it as easy as local computing
 - **Quick physics results** by leveraging distributed resources
 - **Isolate users** from heterogeneity in infrastructure and middleware
 - **Fair sharing** of resources among thousands of physicists
 - **Error handling** fully automated and invisible to users
 - **Workflow auto-adjusts** to reduce delays from service problems
 - **Automatic scheduling among sites** for data processing, physics group activities, and individual physics analysis
 - **Fully integrated** with Data Management System developed at CERN

Using the Full ATLAS Computing Ecosystem

Computing centers distributed worldwide

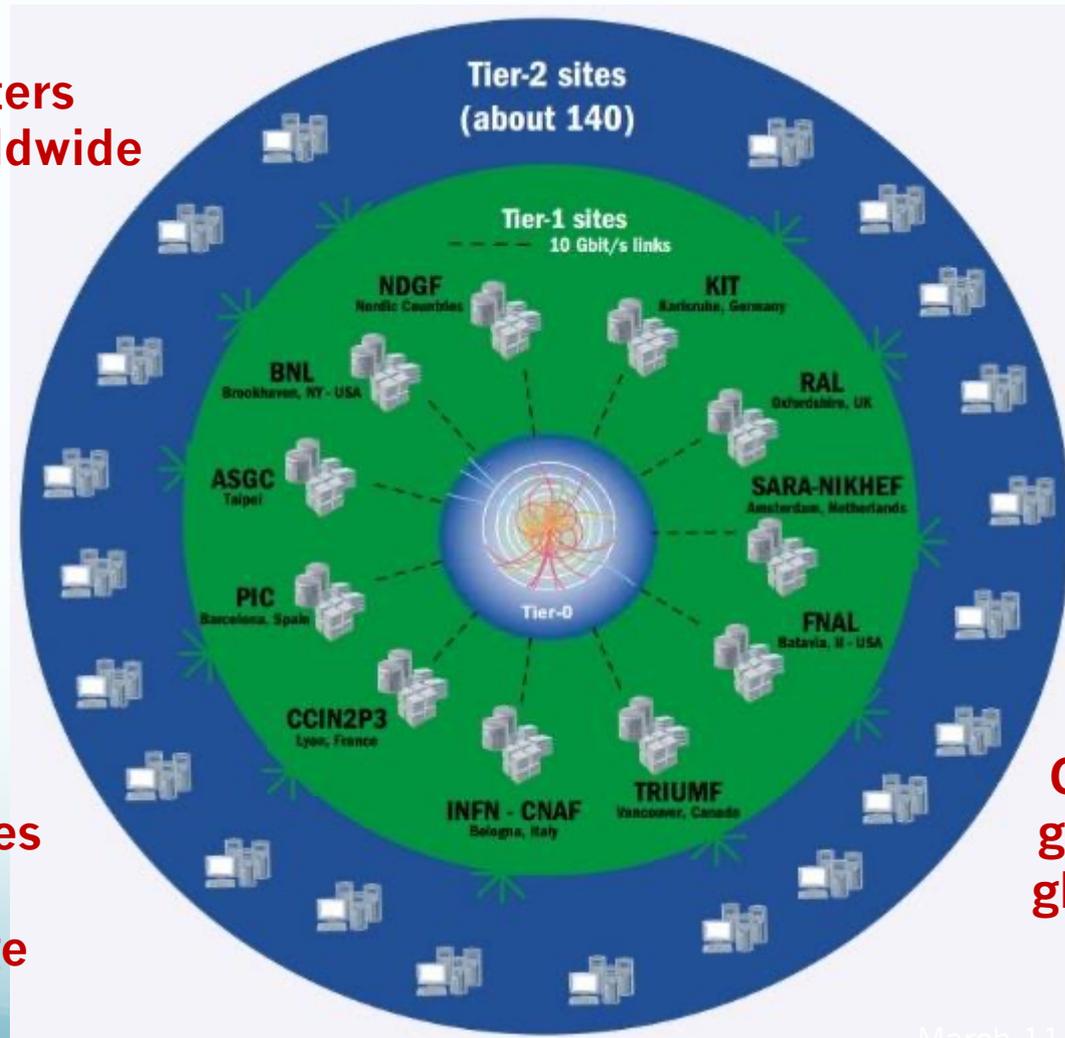
~150 clusters

Heterogeneous

Independently maintained

Provides >800k CPU cores

>500 PB storage



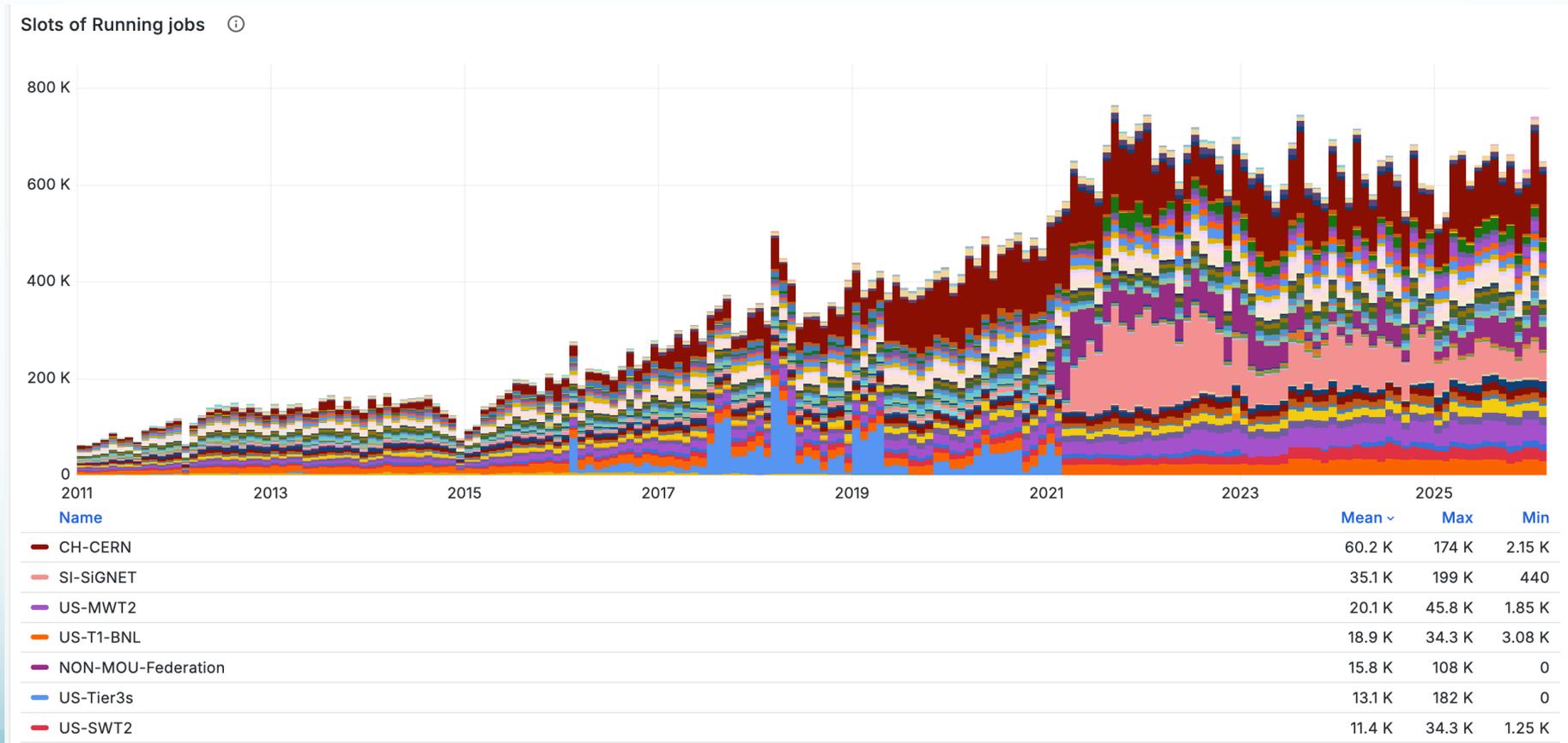
Dozens of applications

Dozens of workflows

Community of thousands of users

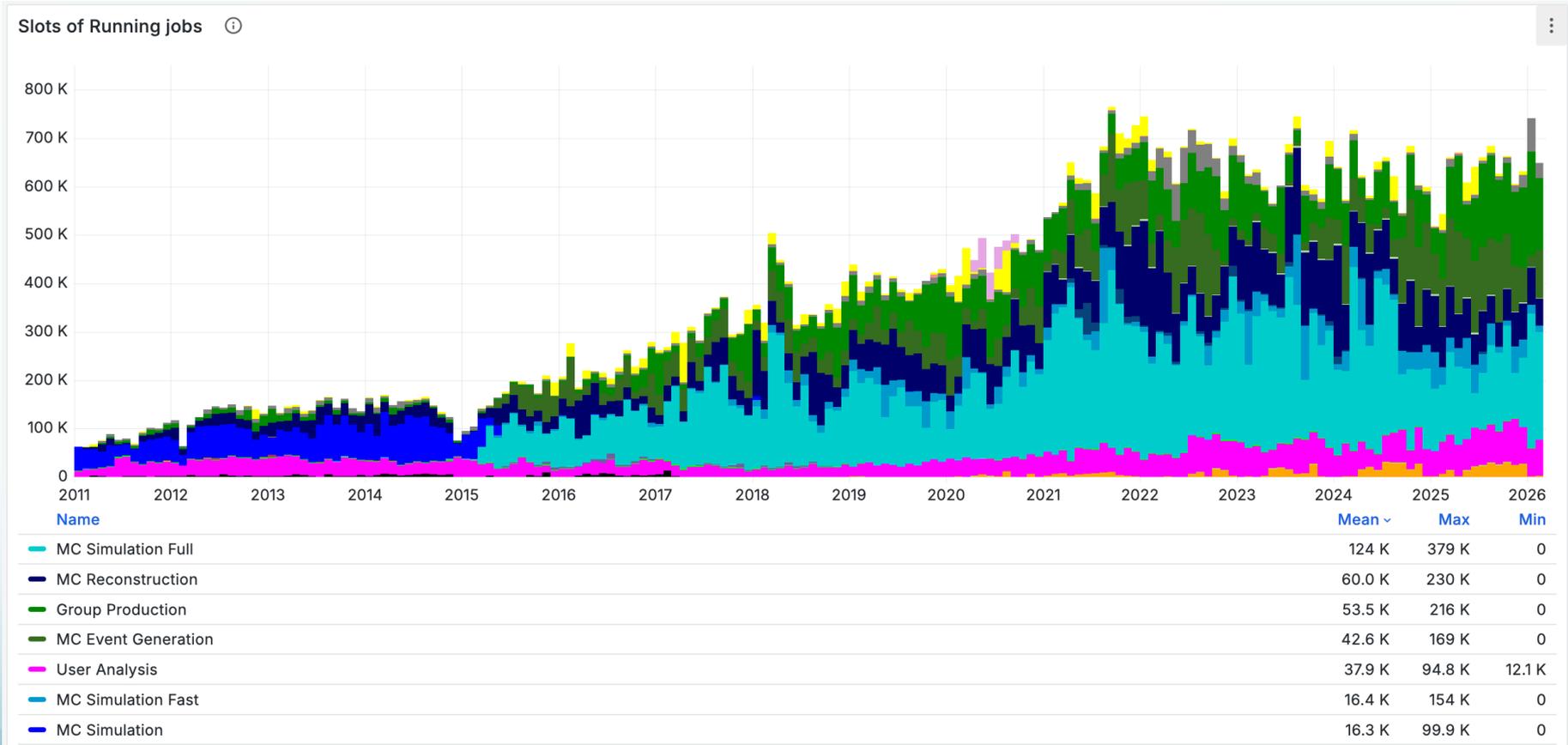
Complex rules:
global fairshare
global priorities
global quotas

Each color represents a distributed resource



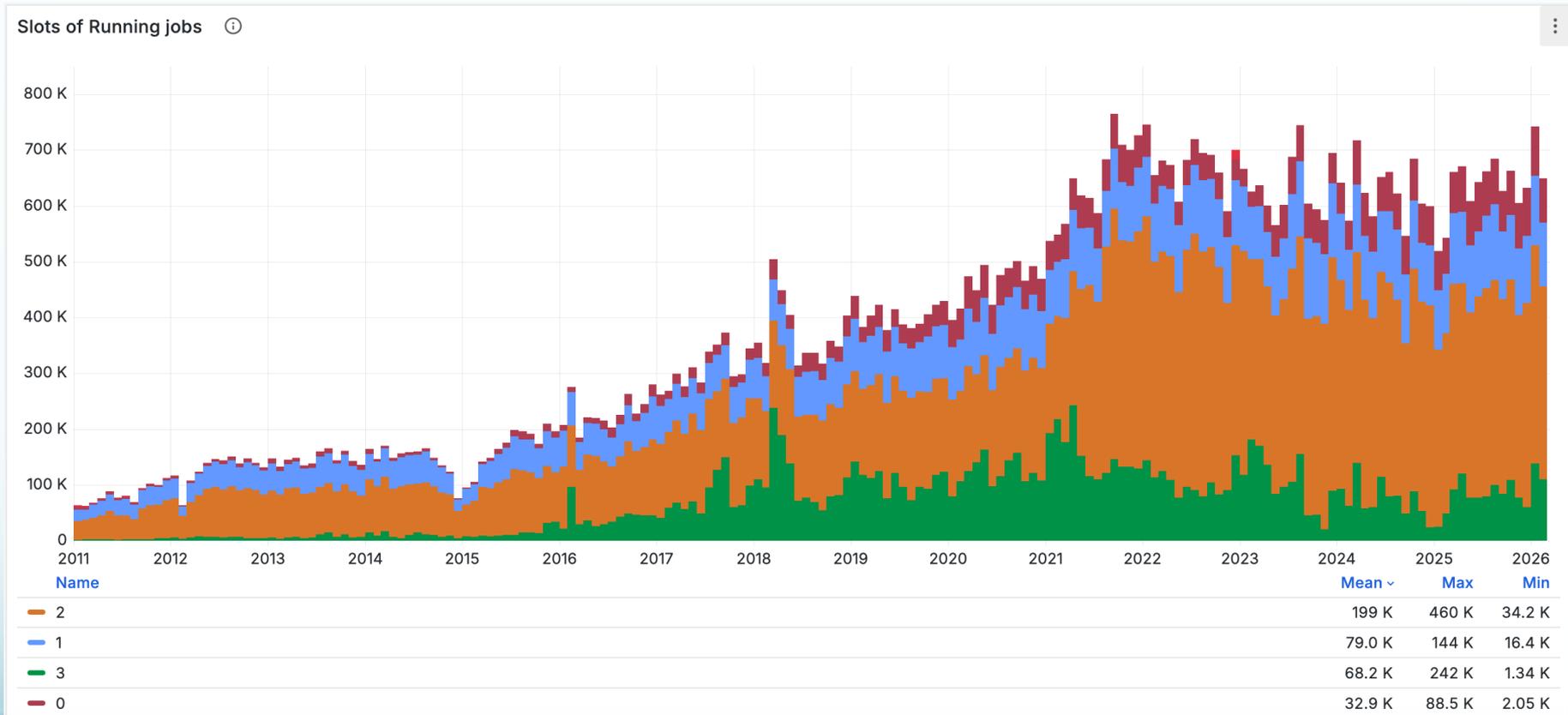
15+ years of continuous monitoring ...

Each color is different work

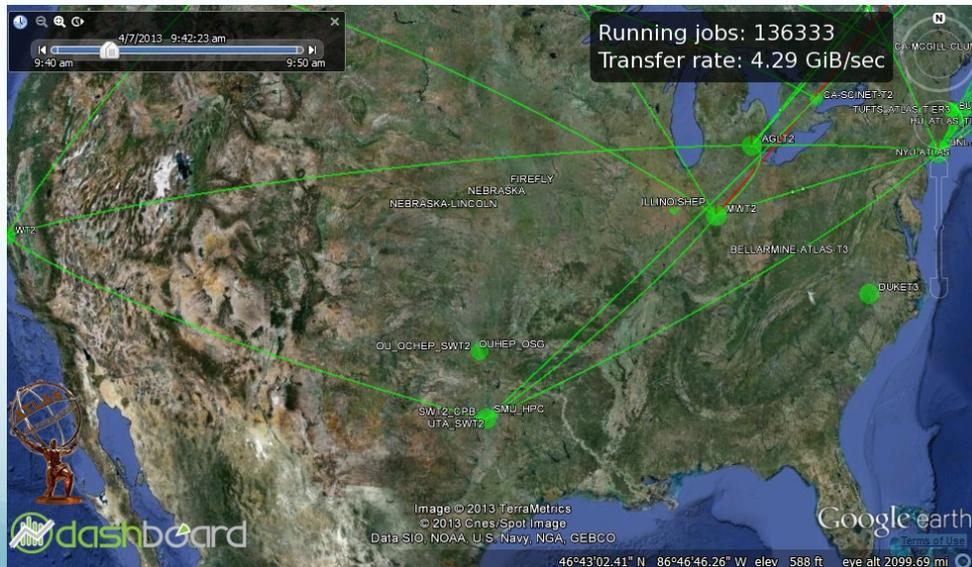
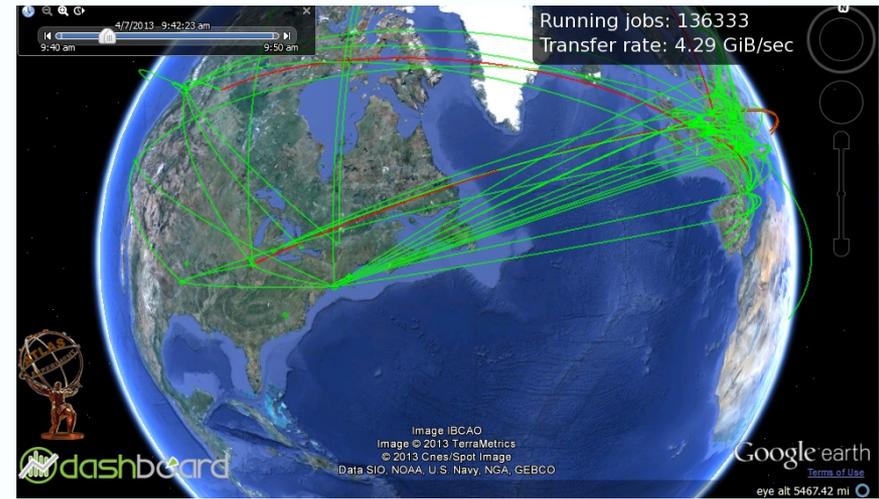
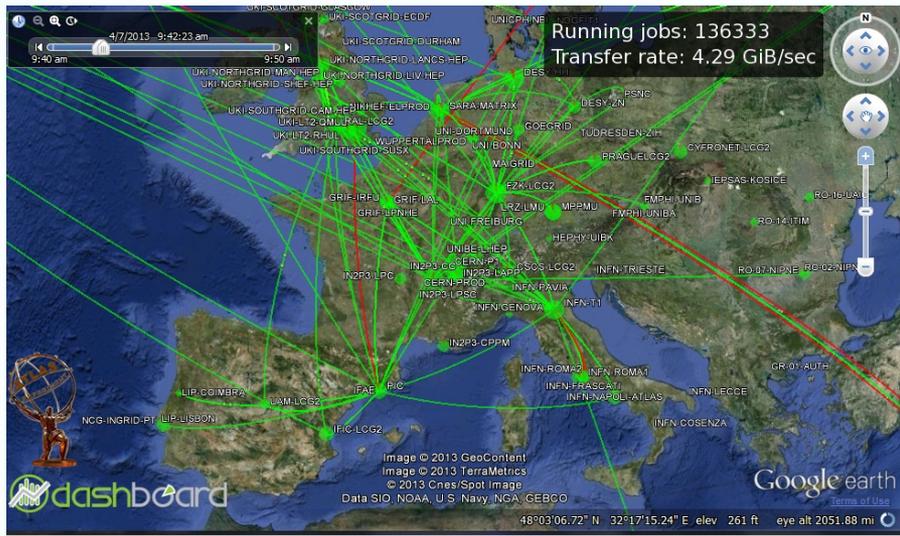


Work is a mix of CPU intensive, data intensive, local or remote access jobs...

Slots available at each Tier



No more hierarchy – inverted Tiers!



- Truly worldwide data movement, storage and processing by physicists

Who are the Users of PanDA in ATLAS

- Everyone in ATLAS!
- Prompt data processing – spillover outside CERN
- Central MC production
 - Managed processing of background samples, signals
 - Multiple processing steps to create final Ntuples
- Data reprocessing
 - Primary data reprocessing
- Derivation production – for improved observables
- Physics Group processing
- User analysis
 - Same engine as production system, but user code
 - Between 1-2 Million jobs submitted daily by users

Paradigm Shift in HEP Computing

- New Ideas from PanDA
 - Distributed resources are seamlessly integrated
 - All users have access to resources worldwide through a single submission system
 - Uniform fair share, priorities and policies allow efficient use of global resources
 - Automation, error handling, and other features in PanDA improve user experience
 - All users have access to same resources
- Old HEP paradigm
 - Distributed resources are independent entities
 - Groups of users utilize specific resources (whether locally or remotely)
 - Fair shares, priorities and policies are managed locally, for each resource
 - Uneven user experience at different sites, based on local support and experience
 - Privileged users have access to special resources

The story of PanDA has parallel in industry – the growth of Cloud Computing

The Technology

- Standards based implementation
 - REST framework – HTTP/S
 - Oracle backend (also support other DB engines)
 - About a dozen Python packages available from GitHub
 - Command-line and GUI/Web interfaces
 - Modular design
 - Modern authentication
 - Use of Open Source components
- Reference
 - <http://news.pandawms.org/>
 - <https://twiki.cern.ch/twiki/bin/view/PanDA/PanDA>

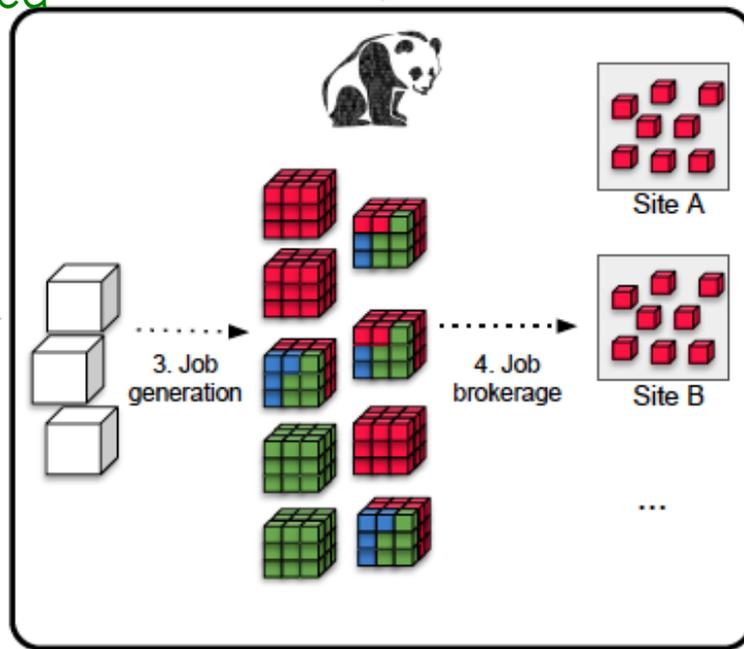
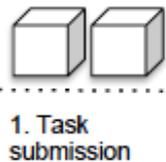
PanDA System for Production and Distributed Analysis

PanDA concepts

ATLAS workload management system developed in 2005



System to manage from 1 job to a million jobs



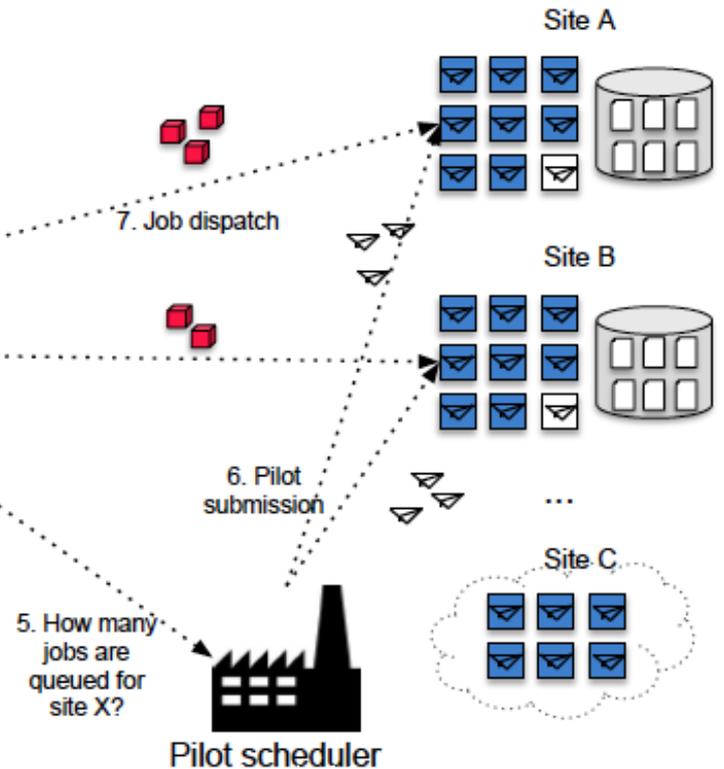
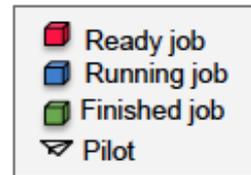
PanDA



Rucio

Global queue

- ↑ 2. Dataset lookup, file locations and replication requests
- ↓



From HTC to HPC to Clouds

- ATLAS computing clusters on the grid are often called HTC – High Throughput Computing
- After highly successful Run 1 using the Grid, we added HPC (eg. supercomputers) and Cloud computing (eg. Amazon and Google) to the mix
- HTC/Grid + Clouds + HPC == Truly Heterogeneous and distributed computing integrated seamlessly through the PanDA ecosystem now used for Run 3
- And now PanDA supports multiple architectures like ARM and GPU in addition to CPU

Lessons Learned after 20+ years

- Large experiments need seamlessly integrated distributed computing systems
 - To leverage heterogenous resources anywhere
 - Must support all workloads and workflows
 - Must support all users with fair-share and priorities
- For ePIC – PanDA will need to grow and evolve
 - Just like for ATLAS experiment over 20+ years
 - PanDA was built to be the heart of a S&C ecosystem for large experiments like ATLAS and ePIC
 - Simple turnkey systems cannot meet future complexity – PanDA is extensible and critical to ATLAS success
- Open to questions...